Tutoriat PS 4: Lab Exercises Explained (Revision on Everything So Far)

Mihai Duzi Lucan Cristian Luminaru Ionuț

1 Laboratory 5

Note: Refer to *Tutoriat 1* for the first two exercises. Variations of those were already practiced.

Exercise 3 – Subpoint 1

Reference: For additional examples and theory, see *Tutoriat 2*, Section 2.2 (Binomial Distributions).

Problem Statement

The probability of meeting someone I know on the subway on my way to university is $p = 9\% = 0.09$.

In one month (30 days), which is more likely:

- not meeting anyone I know on the subway at all, or
- meeting someone exactly 5 times?

Solution

Let X be the number of days (out of 30) when I meet someone I know.

$$X \sim \text{Binomial}(n = 30, p = 0.09)$$

The probability mass function of a Binomial random variable is:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Case 1: No encounters in a month ($X = 0$)

$$P(X = 0) = \binom{30}{0} (0.09)^0 (0.91)^{30} = (0.91)^{30} \approx 0.057$$

Case 2: Exactly 5 encounters in a month ($X = 5$)

$$P(X = 5) = \binom{30}{5}(0.09)^5(0.91)^{25}$$

$$\binom{30}{5} = 142506$$

$$P(X = 5) = 142506 \times (0.09)^5 \times (0.91)^{25} \approx 0.184$$

Comparison:

$$P(X = 0) \approx 0.057 \quad \text{vs.} \quad P(X = 5) \approx 0.184$$

Therefore, it is **more likely** that there will be exactly 5 days in the month when I meet someone I know, rather than none at all.

—

2 Laboratory 4

Exercise 1

Reference: For related examples, see *Tutoriat 1*, Section 3.4.

Problem Statement

A Chinese manufacturer releases a rapid test for detecting COVID-19 infection, with the following characteristics:

- **Sensitivity:** 98% — the probability that the test correctly identifies infected individuals.
- **Specificity:** 99% — meaning the probability that the test is positive when the person is *not* infected is 1%.

Medical studies show that the incidence of COVID in Bucharest is 1 in 1000 people, i.e. $P(S) = 0.001$.

We are asked to compute:

1. If one test comes out positive, what is the probability that I am actually infected?
2. If I take two independent tests and both come out positive, what is the probability that I am actually infected?

—

Notation

S: Person is sick (infected)
H: Person is healthy (not infected)
$+$: Test is positive
$-$: Test is negative

$$P(S) = 0.001, \quad P(H) = 0.999$$
$$P(+|S) = 0.98, \quad P(+|H) = 0.01$$

—

Question 1: $P(S|+)$

Using Bayes' theorem:

$$P(S|+) = \frac{P(+|S) \cdot P(S)}{P(+)}$$

We compute $P(+)$ using the law of total probability:

$$P(+) = P(+|S) \cdot P(S) + P(+|H) \cdot P(H)$$

$$P(+) = 0.98(0.001) + 0.01(0.999) = 0.00098 + 0.00999 = 0.01097$$

Now:

$$P(S|+) = \frac{0.98 \times 0.001}{0.01097} = \frac{0.00098}{0.01097} \approx 0.0893$$

$$\boxed{P(S|+) \approx 8.93\%}$$

—

Question 2: $P(S|++)$

Assuming the two tests are independent:

$$P(++|S) = (P(+|S))^2 = (0.98)^2 = 0.9604$$
$$P(++|H) = (P(+|H))^2 = (0.01)^2 = 0.0001$$

Now:

$$P(++) = P(++|S) \cdot P(S) + P(++|H) \cdot P(H)$$

$$P(++) = 0.9604(0.001) + 0.0001(0.999) = 0.0009604 + 0.0000999 = 0.0010603$$

Finally:

$$P(S|++) = \frac{P(++|S) \cdot P(S)}{P(++)} = \frac{0.9604 \times 0.001}{0.0010603} \approx 0.906$$

$$\boxed{P(S|++) \approx 90.6\%}$$

—

Exercise 2

Problem Statement

In a television game show, a contestant is presented with three identical doors. Behind one door there is a car, and behind the other two there are goats. The contestant picks one of the doors; then the host (who knows the locations of the car and the goats) opens one of the two remaining doors, revealing a goat. The host then asks the contestant whether they want to keep their original choice or switch to the other unopened door. Is it advantageous for the contestant to switch? Perform a numerical simulation to estimate the probability of winning using each of the two possible strategies (always switch vs. always stay).

Question 1: Probability of winning by switching

Note: It is sufficient to condition on the contestant having initially chosen door 1; by symmetry the same result holds if the contestant initially chose door 2 or door 3, and the overall probability is the average over the three initial choices.

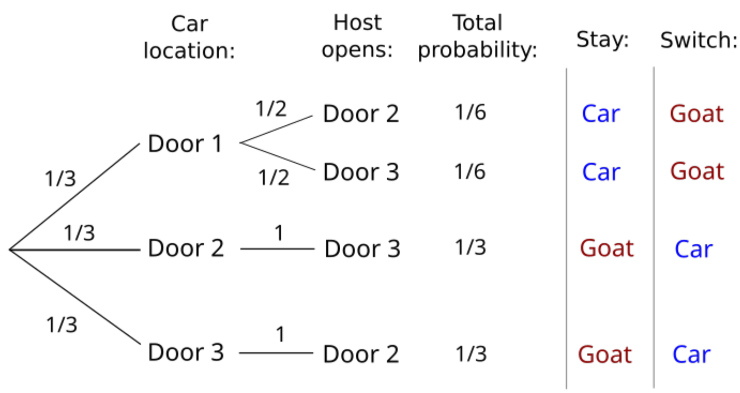

Figure 1: Monty Hall decision tree when the contestant initially chooses door 1

By definition, the conditional probability of winning by switching, given that the contestant initially chooses door 1 and the host opens door 3, is the probability of the event “car is behind door 2 and host opens door 3” divided by the probability of the event “host opens door 3”.

Define events:

$$\begin{aligned} C_i &: \text{ the car is behind door } i, \\ H_j &: \text{ the host opens door } j, \\ I_k &: \text{ the contestant initially chooses door } k. \end{aligned}$$

We want

$$P(\text{win by switching} \mid I_1, H_3) = P(C_2 \mid I_1, H_3) = \frac{P(C_2, H_3 \mid I_1)}{P(H_3 \mid I_1)}.$$

Compute the numerator $P(C_2, H_3 \mid I_1)$. Condition on where the car is:

$$P(C_2, H_3 \mid I_1) = P(C_2 \mid I_1)\, P(H_3 \mid C_2, I_1).$$

Since the car is equally likely behind any door a priori, $P(C_2 \mid I_1) = \frac{1}{3}$. If the car is behind door 2 and the contestant picked door 1, the host *must* open door 3 (there is only a goat behind door 3), hence $P(H_3 \mid C_2, I_1) = 1$. Therefore

$$P(C_2, H_3 \mid I_1) = \frac{1}{3} \cdot 1 = \frac{1}{3}.$$

Compute the denominator $P(H_3 \mid I_1)$ by total probability over the car location:

$$\begin{aligned} P(H_3 \mid I_1) &= \sum_{i=1}^{3} P(C_i \mid I_1)\, P(H_3 \mid C_i, I_1) \\ &= P(C_1 \mid I_1)P(H_3 \mid C_1, I_1) + P(C_2 \mid I_1)P(H_3 \mid C_2, I_1) + P(C_3 \mid I_1)P(H_3 \mid C_3, I_1). \end{aligned}$$

Evaluate each term:

- If the car is behind door 1 (probability 1/3), the host chooses uniformly between doors 2 and 3 (both goats), so $P(H_3 \mid C_1, I_1) = \frac{1}{2}$.
- If the car is behind door 2 (probability 1/3), the host must open door 3, so $P(H_3 \mid C_2, I_1) = 1$.
- If the car is behind door 3 (probability 1/3), the host cannot open door 3 (it has the car), so $P(H_3 \mid C_3, I_1) = 0$.

Thus

$$P(H_3 \mid I_1) = \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0 = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}.$$

Putting numerator and denominator together:

$$P(C_2 \mid I_1, H_3) = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.$$

Therefore, $\boxed{P(\text{win by switching} \mid I_1, H_3) = \frac{2}{3}}$. By symmetry (and averaging over the three possible initial choices), the unconditional probability of winning by always switching is also 2/3.

Question 2: Probability of winning by staying

The probability of winning by staying (i.e. keeping the initial choice) given I_1 and H_3 is

$$P(\text{win by staying} \mid I_1, H_3) = P(C_1 \mid I_1, H_3) = \frac{P(C_1, H_3 \mid I_1)}{P(H_3 \mid I_1)}.$$

We already computed $P(H_3 \mid I_1) = \frac{1}{2}$. Moreover,

$$P(C_1, H_3 \mid I_1) = P(C_1 \mid I_1)\, P(H_3 \mid C_1, I_1) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6},$$

so

$$P(C_1 \mid I_1, H_3) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}.$$

Thus $\boxed{P(\text{win by staying} \mid I_1, H_3) = \frac{1}{3}}$. This matches the well-known result that always staying yields win probability 1/3, while always switching yields 2/3.